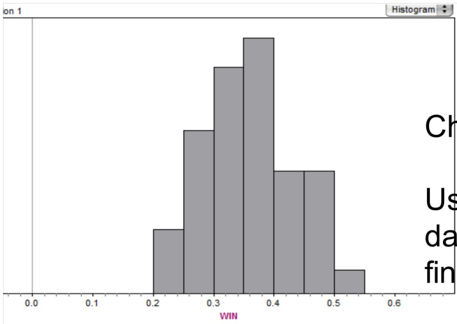


Ch. 19 Activity:

Using class data given, finish up to #11

Then get out Ch. 19 notes

Mean = 0.346
s = 0.07

CH. 19: Confidence Intervals about a proportion

From Ch. 18....Distirbution of sample proportions ($\hat{p}$) follow the model (if checks pass):

$$N\left(p, \sqrt{\frac{pq}{n}}\right)$$

However, most of the time we don't know...
the population proportion (p)

* We take samples and calculate __$\hat{P}$__ to try to find __P___

**** Since we don't know p, we can't find std. deviation*** $\sqrt{\frac{pq}{n}}$

* __$\hat{P}$___ is the **estimate** for __P____
(estimator)

* So...we can estimate the std. dev. with STANDARD ERROR:

$$SE_{\hat{p}} = \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

CONFIDENCE INTERVALS... Intro

EXAMPLE: It is election season. You open the newspaper and see a headline: ***57% of the nation favors the Democratic candidate for President.***

a) Does EXACTLY 57% of the nation favor the Democrat? Why?

NO. Sample

b) As you keep reading, you see the following: ***There is a margin of error of 5%***. What does this mean about the TRUE PERCENT of the nation that favors the Democrat? Do you think that the pollsters have a really good idea of how many people will vote for the Democrat on Election Day?

57% ± 5%

(52%, 62%)

c) As you keep reading, you see the following: ***The results of this poll are given with 70% confidence.*** Do you think that their results are reliable? Why?

NO

d) What if they change their confidence to 95%? Do you think that their results are reliable? Why?

Yes

e) What if the results are as follows: 57% will vote Democrat, margin of error of 15%, 99% confidence. What can you say about the results? Are they reliable? Why?

57% ± 15%

(42%, 72%)

CONFIDENCE INTERVAL:

* Based on $\hat{p}$, and the sampling distribution of $\hat{p}$

* Start with $\hat{p}$, give ourselves a margin of error on either side

** MOE = b/c many of our samples are not perfect*

* Size of interval is based on sample size and level of confidence:

* Larger sample size = smaller interval (more accurate)
* Larger confidence level = larger (wider) interval
(need more room for error)

Confidence interval: Basic Setup:

ESTIMATE ± MOE = (a, b)

For a one-proportion sample:

$$\hat{p} \pm \boxed{(Z^*)(SE_{\hat{p}})} = \hat{p} \pm (Z^*)\sqrt{\frac{\hat{p}\hat{q}}{n}} = (a, b)$$

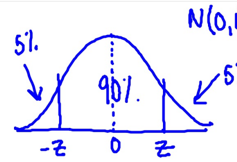
Z* = critical value

= ___ % of the data is between ± Z (confidence level)

* Common Critical Values:

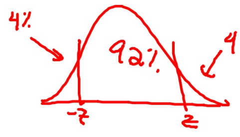
Level of Conf.	Z*
90%	1.645
95%	1.960
99%	2.576

N(0,1)

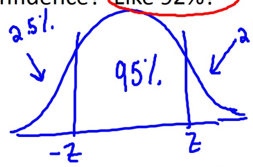


invnorm(0.05, 0, 1)

What about another leve of confidence? Like 92%?



Z = 1.751



invnorm(0.025, 0, 1)

Whenever we create a confidence interval, we write a sentence interpretation:

We are _95_ % confident that the true % of _______ is between _a_ and _b_ %.

*must write every time

Example: President Obama 45% approval rating, MOE of 3%, 95% confidence:

We are 95% confident that the true % approval rating of President Obama is between 42% and 48%.

0.42 0.48

Conditions (for 1-Prop Z-Intervals)

STATE

(1) Simple Random Sample (SRS)

(2) $n\hat{p} \geq 10$
$n\hat{q} \geq 10$

** must use $\hat{p}$ because we don't know p

(3) population ≥ 10*n*

If all conditions are met, then we can say $\hat{p}$ has model:

conditions met, $N\left(\hat{p}, \sqrt{\frac{\hat{p}\hat{q}}{n}}\right)$, 1 prop. Z-Interva

and we can use normal model for 1-Prop Z-Interval

Example: We want to know the real improvement rate for a new medication. We conduct an experiment and find that out of 53 subjects, 27% of them report improvement with the new medications. Create a 95% confidence interval (and interpret)

$n = 53$
$\hat{p} = 0.27$
$C = 95\%$

Conditions
1) SRS
2) $n\hat{p} \geq 10$
$n\hat{q}$
3) pop $\geq 10n$

1) assumed subjects are representative
2) $(53)(0.27)$
$(53)(0.73) \geq 10$
3) there are more than 530 people taking med.

Cond. met, $N\left(0.27, \sqrt{\frac{(0.27)(0.73)}{53}}\right)$, 1 prop. Z-Interval

0.061

$$\hat{p} \pm (Z^*)\left(\sqrt{\frac{\hat{p}\hat{q}}{n}}\right)$$

$$(0.27) \pm (1.96)\sqrt{\frac{(0.27)(0.73)}{53}}$$

0.27 ± 0.11956

$(0.15044, 0.38956)$

We are 95% confident that the true % of improvement w/ the medication is btw. 15.044% and 38.956%.

Example: We take a simple random sample of 95 Bucks county residents and find that only 20 of them approve of a new property tax to pay for repairs to local roads. Estimate with 99% confidence the true percent of people who approve of the tax.

$n = 95$

$\hat{p} = 20/95 = 0.211$

$C = 99\%$

Example: We take a simple random sample of 95 Bucks county residents and find that only 20 of them approve of a new property tax to pay for repairs to local roads. Estimate with 99% confidence the true percent of people who approve of the tax.

$n = 95$

$\hat{p} = 20/95 = 0.211$

$C = 99\%$

STATE	CHECK
- SRS	- stated SRS
- $n\hat{p} \geq 10$	- $(95)(0.211) \geq 10$
$n\hat{q}$	$(95)(0.789)$
- pop $\geq 10(n)$	- all bucks residents ≥ 950

conditions met --> N(0.211, 0.0419) --> 1 prop Z Interval

$\hat{p} \pm (Z^*)\left(\sqrt{\frac{\hat{p}\hat{q}}{n}}\right) = (0.211) \pm (2.576)(0.0419) = (0.1031, 0.3189)$

We are 99% confident that the true % of bucks county residents who approve the tax is btw. 10.31% and 31.89%.

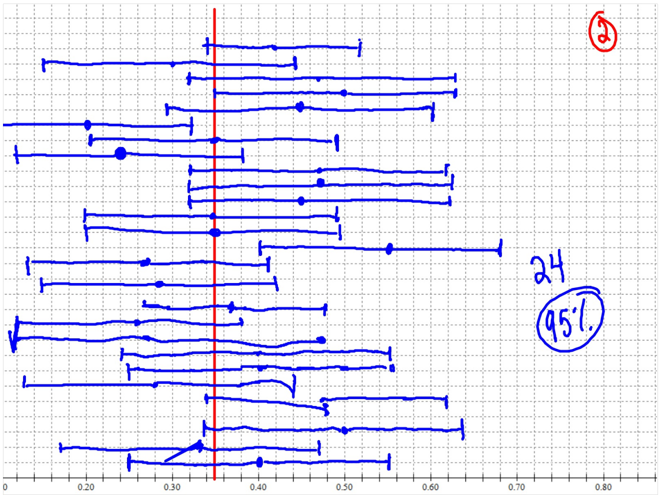
Go back and finish the Ch. 19 activity

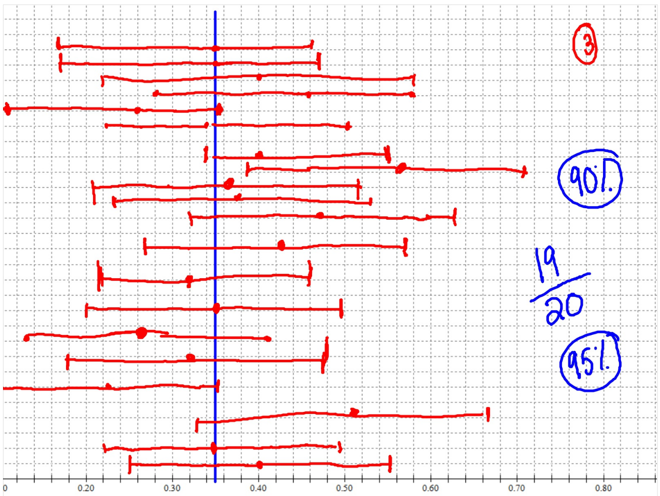
#12

$n = 40$

$\hat{p} =$

(a, b)





What does 95% confidence really mean??

In repeated samples of size __n__, the conf. interval created will catch the true parameter (p) 95% of the time.

95% of samples of this size __n__ will produce confidence intervals that catch the true parameter.

So for our activity:

In repeated samples of size ~~30~~ 40, the conf. interval created will catch the true winning percent of the GAME 95% of the time.

Things that affect the MOE:

* sample size

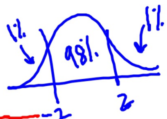
* confidence level

$\hat{p} \pm \boxed{z^* \sqrt{\frac{\hat{p}\hat{q}}{n}}} = (a, b)$ = MOE

If we want a particular MOE, we can set a level of confidence and a sample size in order to attain that MOE.

Example: Let's go back to our example about the improvement rate with a new medication. We found 27% improvement. How many subjects would we need in a new experiment to make a 98% conf. interval while still keeping a 5% MOE?

$m = z^* \sqrt{\frac{\hat{p}\hat{q}}{n}}$



$0.05 = (2.326)\sqrt{\frac{(0.27)(0.73)}{n}}$

$n = 426.5461....$

427 people

always round up

$\frac{0.05}{2.326} = \frac{2.326 \sqrt{\frac{(0.27)(0.73)}{n}}}{2.326}$

$0.0214....^2 = \left(\sqrt{\frac{(0.27)(0.73)}{n}}\right)^2$

$4.6208 \times 10^{-4} = \frac{(0.27)(0.73)}{n}$

$(0.27 \cdot 0.73)/Ans$

NOTE: If you are not given p, you can use **0.50**.

Why?

* It doesnt favor either outcome (success or failure)

* It will give you the largest sample size. If you try any other value of p, you will notice a smaller numerator.

Ex: (0.5)(0.5) = 0.25
(0.6)(0.4) = 0.24
(0.7)(0.3) = 0.21

$z^* \sqrt{\frac{\hat{p}\hat{q}}{n}}$ $n = 400$

Example: What sample size must be used to estimate the outcome of a political election with a margin of error of 3% and 99% confidence?

$0.03 = (2.576)\sqrt{\frac{(0.5)(0.5)}{n}}$

$n = 1843.2711....$

$n = 1844$ people

$\hat{p} = 0.30$

Example: What sample size must be used to estimate the true percent of left-handed people in the nation with 90% confidence and a margin of error of 8%? Assume that it has been shown in previous research that the percent of left-handed people was 38%.

$C = 90\%$
$m = 8\%$
$\hat{p} = 0.38$

$0.08 = 1.645\sqrt{\frac{(0.38)(0.62)}{n}}$

$\boxed{n = 100 \text{ people}}$

Confidence intervals- Formula sheet:

III. Inferential Statistics

Standardized test statistic: $\dfrac{\text{statistic} - \text{parameter}}{\text{standard deviation of statistic}}$

Confidence interval: statistic $\pm$ (critical value) $\cdot$ (standard deviation of statistic)

$\hat{p} \pm (z^*)(\sqrt{\ \ })$

Single-Sample

Statistic	Standard Deviation of Statistic
Sample Mean	$\frac{\sigma}{\sqrt{n}}$
Sample Proportion	$\sqrt{\frac{p(1-p)}{n}}$

$\bar{x} = 75$
$n = 100$
95% conf. int.

$\bar{x} \pm (\ \ )(\sigma/\sqrt{n})$
$\rightarrow (a, b)$

Two-Sample

Statistic	Standard Deviation of Statistic
Difference of sample means	$\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$ Special case when

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

We take a survey and find that 23 out of 100 people said "YES" to our survey. Create a 96% confidence interval.

CALCULATOR: (See p. 448 in book)

$\hat{p} = 23/100 = 0.23$
$n = 100$
$C = 96\%$

① Conditions

② Cond. met, $N\left(0.23, \sqrt{\frac{(0.23)(0.77)}{100}}\right)$

1 prop. Z-Interval

③ $0.23 \pm (2.054)\sqrt{\frac{(0.23)(0.77)}{100}} = (0.14357, 0.31643)$

④ We are 96% conf that.....

$\hat{p} = 0.84$

$n = 300$

$\frac{x}{n} = \frac{252}{300}$ $\quad \sqrt{\frac{\hat{p}\hat{q}}{n}}$

PRACTICE:

p. 456 #17 & 27

b) We are 95% conf. that....

c) In repeated samples....

17) p-hat = 91/582 = 0.1564   n= 582

STATE	CHECK
- SRS	stated random sample
- $n\hat{p} \geq 10$	$91 \geq 10$
$n\hat{q} \geq 10$	$491 \geq 10$
- pop $\geq 10n$	There are more than 5820 accidents

conditions met --> N(0.1564, 0.0151) --> 1 prop Z Interval

$0.1564 \pm (1.96)\sqrt{\frac{0.1564 \cdot 0.8436}{582}} = (0.12685, 0.18586)$

(b) We are 95% confident that the true % of accidents with teenage drivers is between 12.685% and 18.586%.

(c) In repeated samples of size 582, the intervals created will catch the true percent of accidents with teenage drivers 95% of the time.

(d) 1 in 5 = 1/5 = 20%

Our interval contradicts this statement. We are 95% confident that the true % of accidents with teenage drivers is LESS than 20%.

27) p-hat = 32/153 = 0.2092 n = 153

(a) STATE	CHECK
- SRS	assumed representative
- $n\hat{p} \geq 10$	$32 \geq 10$
$n\hat{q} \geq 10$	$121 \geq 10$
- pop $\geq 10n$	There are more than 1530 deer

conditions met --> N(0.2092, 0.0329) --> 1 prop Z Interval

$$0.2092 \pm (1.645)\left(\sqrt{\frac{0.2092*0.7908}{153}}\right) = (0.15507, 0.26323)$$

We are 90% confident that the true % of deer with lyme disease is between 15.507% and 26.323%.

(b) original MOE = 0.05403

$$\text{half} = 0.027015 = (1.645)\sqrt{\frac{(0.2092)(0.7908)}{n}}$$

n = 614

$$\sqrt{\frac{\hat{p}\hat{q}}{n}} = \frac{\sqrt{\hat{p}\hat{q}}}{\sqrt{n}} \cdot \frac{1}{2}$$

or

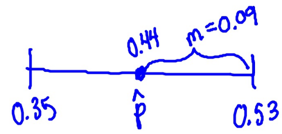
to cut a MOE in half, you need to multiply your sample size by 4.

n = 153 * 4 = 612

1/3

Confidence interval (0.35, 0.53)

(a) find the sample proportion



(b) find the margin of error

(c) Find the level of confidence if the sample size is 200.

HW: p. 455 #8, 22, 24, 29, 30